AD-A147 523

FINAL REPORT

for

NEW TECHNOLOGY HIGH EFFICIENCY AND HIGH STABILITY COLOR MOTION VIDEO COMPRESSION SYSTEM

submitted by

AVELEX

310 Bonifant Road
Silver Spring, MD 20904

August 15, 1984

Contract N00014- 84-C-0676

SUMMARY

New video signal compression technology including a novel mathematical transform was previously developed by Avelex to provide for low bit rate transmission of full motion monochrome imagery. The nature of this new technology has suggested that chrominance might be added to provide full color imagery with a rather small increase in the bit rate required for transmission above that required for the monochrome transmission alone. The subject task was proposed to DARPA to demonstrate the feasibility of a full color, full motion video compression system which could be used for Defense Department and other Government Agencies, as well as commercial applications. Two signal processing techniques were proposed for the DARPA work which have been implemented and experimentally evaluated with successful outcomes, and the results herein reported. Additionally, a Videotape showing the experimental results is submitted as auxilliary information to this final report.

BACKGROUND

A combination of Intra-frame and Inter-frame methods for compression of full motion monochrome imagery has been developed by Avelex and others, albeit based on somewhat different techniques. Avelex previously developed the novel Triangle and Pyramid mathematical transform methods, Ref 1, for compression of audio and video signals as well as a group of methods for Inter-frame compression of video signals. Principal image performance break-throughs compared with any prior technology are three-fold. First, the new transform does not require a sub-dividing or "blocking" process of the image and hence does not generate artifacts relating to blocking which appear as mosaics or tiles superimposed on the image. Second, the interframe methods prevent any remnants from previous images in a sequence of images producing motion imagery from appearing after they are no longer valid, regardless of the transmission channel capacity of the link being employed or the amount of motion in the image. Third, relative motion between the camera and an object does not result in image break-up, as happens with most other high compression video codecs. This latter capability permits use of the codec in difficult situations where relative motion between camera and object can exist. To accomplish this performance within the limitations of a fixed rate transmission channel, edges of objects in motion may temporarily lose some definition while they remain in motion. The loss of resolution condition depends upon the amount of motion presented to the compression equipment relative to the capacity of the transmission channel being employed. Under many conditions the amount of image motion may not be enough to result in loss of any resolution.

As well as performance advantages, the Pyramid Transform provides economic benefits. Hardware required to perform the Pyramid mathematical transform is substantially reduced from that necessary to perform the often employed Cosine Transform. The Pyramid Transform can be calculated using a "fast" calculation method in which only additions, subtractions and binary shifts are required whereas the Cosine Transform requires non-trivial multiplications and complex arithmetic. Despite the simplicity of Pyramid Transform calculations, the basis functions are all smooth and do not result in artifact generation when higher frequency coefficients are omitted or coefficients more roughly quantized.

Also, the Pyramid Transform employs basis functions which are naturally finite on the original two dimensional signal space, which results both in a sparse calculation matrix and not having to sub-divide or "block" the image into many mosaics prior to performing the transform operation on each resulting block individually. By not having to "block" the image no "tiling" effects are ever observable in a Pyramid Transform reconstructed image.

It has been known since the work toward the adoption of the N.T.S.C. color television standard in the U.S. in 1954 that the human visual acuity to chrominance portions of an image, as opposed to luminance portions, is considerably less than required for a purely black and white video transmission. This fact led to the significant reduction in bandwidth required for transmission of chrominance portions of video and, along with the development of a novel frequency interleaving technique, the ability to transmit both monochrome and chrominance portions of an image in the same bandwidth as previously required for monochrome transmission alone.

Also, the N.T.S.C. technique inherently relies on the fact that the chrominance portions of the image have a high degree of spatial correlation with their monochrome counterparts, at least for television receivers which do not completely separate the chrominance subcarrier from the directly displayed monochrome component prior to display. Difficulty arises with the basic assumption that a chrominance subcarrier will be invisible to a human observer. Although the subcarrier frequency has been selected to be an odd multiple of one-half the line rate and should, with the aid of persistence of vision, add to the monochrome component in one frame and subtract from it during the next resulting in visual subcarrier cancellation, the non-linear characteristic of the picture tube employed causes incomplete cancellation of the resulting brightness of the two frames. The presence of the subcarrier added to the monochrome video signal and directly applied to the luminance of the video display device actually increases the output brightness of the image to a degree proportional to the saturation of the color (amplitude of the subcarrier) in a particular area. If dramatic color saturation

changes were to occur in an image area wherein the brightness component remained constant, the perceived brightness would actually change. However, the natural correlation between chrominance changes and luminance changes results in the aforementioned effect not being troublesome in practice. An example of the effect of the monochrome signal from a first video signal and the chrominance signal from a second source transmitted as a single encoded signal is given in Ref 2, page 213. When received and displayed on a monochrome receiver the effect of the brightness alteration due to the chrominance signal is very apparent. However, when the monochrome and chrominance signals are both taken from the same source, any resulting distortion of the luminance signal is not apparent.

The first of the two aspects of the work of this project to be demonstrated relies on the lower human visual acuity to chrominance image portions relative to monochrome portions, to effect an efficient transmission of color image components. The second aspect takes advantage of the usually natural alignment of image transitions (edges) between chrominance and monochrome components of the same image to more efficiently encode chrominance components for transmission. This latter efficiency results from using the same, or nearly the same, overhead map which signals the receiver where to utilize non-zero value monochrome transform coefficients to also signal the usage of the transmitted non-zero value chrominance coefficients.

The motivation for the belief that these aspects of the task would work are as follows. The Pyramid Transform has a strong frequency characteristic and consequently the resulting transform coefficients represent specific spatial frequencies in an image. Therefore, if human visual spatial acuity is frequency limited to a predetermined degree, it should be possible to omit from transmission transform coefficients representing spatial frequencies above those which are observable. The receiver, as a result of this process, will be pre-programmed to assume that the non-transmitted higher spatial frequencies of the chrominance components have a value of zero and accordingly reconstruct a chrominance image component with less detail. Secondly, the fact that chrominance edges are primarily co-located with monochrome edges in naturally occurring images leads to the presumption that non-zero value transform coefficients for chrominance components for the Pyramid Transform occur in the same locations as their monochrome counterparts. This latter presumption arises because the Pyramid Transform yields zero-value transform coefficients, which are not transmitted, in all smooth image areas and non-zero value coefficients, which are transmitted, only in areas which possess changes or edges. In this regard, smooth is defined as meaning any image area wherein the monochrome or chrominance image component has a constant or spatially linearly varying characteristic.

THE TRIANGLE AND PYRAMID TRANSFORMS

The Triangle and Pyramid transforms are respectively one and two dimensional signal transforms which are fully described in Ref 1. Of first interest for this task are the Fourier Transforms of the basis functions of the Triangle and Pyramid transforms so as to determine what frequency groups are carried by which transform coefficients. This is of direct use in determining which Pyramid transform coefficients are required to obtain a certain desired spatial frequency response for chrominance image components.

The one dimensional Triangle Transform is most readily discussed without the complication of the second spatial dimension. The Pyramid Transform involves almost identical application of two Triangle Transforms, one in a direction orthogonal to the other, a process usually employed in other two dimensional transformations. Thus a basis function having a triangular shape in one dimension has a pyramid shape in two dimensions.

The Triangle Transform is organized into a number of bands, Bands 1 through N, each having coefficients, and the Band 1 also having so-called "B" functions. The value of N can vary over a rather wide range, but has been selected as five for the present discussion. In the forward transform direction a decimation is performed using triangular weighting functions and coefficient weighting functions shown in Figure 1. The coefficients have the useful property that input samples which form a straight line over the span of the coefficients' footprint result in zero values, which usually do not require transmission in a video compression system. This property is extremely advantageous in providing for operation of the second part of this contract work to be discussed later. The coefficients resulting from the first decimation are called the Band N coefficients.

For an input signal having P samples there are P/2 triangular function outputs, also called "B" functions, and P/2 coefficients in Band N. For a two dimensional input signal there are P/4 "B" functions and 3*P/4 coefficients. The coefficients receive no further transform processing whereas the "B" functions provide the signal inputs to the Band N-1 processing.

The operation of the Band N-1 processing is identical to that of Band N and results in "B" functions and coefficients. Figure 2 shows the important result that performing two successive triangular decimations is itself a triangular weighting function of the original input samples. The triangular weighting function centered on P3 yields B52; the weighting function centered on P5 yields B53 and so forth. In the next lower band the weighting function which is twice as wide as the Band N function and

centered on B53 yields B41. The equations for the weighting functions are as follows:

$$B52 = 0.25*P2 + 0.5*P3 + 0.25*P4,$$
$$B53 = 0.25*P4 + 0.5*P5 + 0.25*P6, \text{ and}$$
$$B54 = 0.25*P6 + 0.5*P7 + 0.25*P8, \text{ such that,}$$

$$B41 = 0.25*B52 + 0.5*B53 + 0.25*B54, \text{ or,}$$

$$B41 = 0.25*(0.25*P2 + 0.5*P3 + 0.25*P4) + 0.5*(0.25*P4 + 0.5*P5 + 0.25*P6) + 0.25*(0.25*P6 + 0.5*P7 + 0.25*P8).$$

$$B41 = (1/16)*P2 + (2/16)*P3 + (3/16)*P4 + (4/16)*P5 + (3/16)*P6 + (2/16)*P7 + (1/16)*P8.$$

Thus the weighting function in each successive lower band is another triangular function with twice the width and half the height of the preceding higher band. Although all triangular weighting functions below Band N have non-trivial (not a power of the base two) multipliers, the process of progressive decimation achieves the desired lower band weighting functions without performing any non-trivial multiplications.

The frequency response of the triangular function is readily found via the Fourier Integral and an input triangle of height V and width at the base of 2*T to be:

$$F(W) = V * T * (X1*X1), \text{ where } X1 = SIN(W*T/2)/(W*T/2).$$

The familiar Sin(X)/X function has its first zero at X=pi (pi= 3.14159) or equivalently at W = 2*pi/T. Since T has been shown to double in each successive lower Band and since W and T are reciprocally related, the value of W where the frequency response is first zero in each band decreases by a factor of two for each successive lower band.

The frequency response of a five band transform system is shown in Figure 3. From this it can be seen that the difference between the input signal bandwidth and the Band 5 "B" function bandwidth, taken at the 85% point is about one octave. We can deduce therefrom that the Band 5 coefficients occupy primarily the frequency spectrum from one half the maximum to the maximum input signal frequency. By similar reasoning the Band 4 coefficients occupy a spectrum primarily between one fourth and one half of the maximum input frequency.

Applying this to the constructed experimental system wherein the sampling frequency is 9.5 MHz. and the maximum permitted input signal bandwidth is 4.75 MHz. the Triangle Transform coefficients in one dimension are seen to roughly occupy the following frequency regions:

BAND	FREQUENCY RANGE - MHz.
5	2.375 to 4.75
4	1.19 to 2.375
3	0.6 to 1.19
2	0.3 to 0.6

The horizontal bandwidths specified for transmission of chrominance components for the commercial color system in the United States are:

I Axis: 0 to 1.3 MHz.
Q Axis: 0 to 0.5 Mhz.

The I axis corresponds to the axis of maximum visual acuity and occurs for colors in the red and orange regions. The Q axis corresponds to the axis of minimum visual acuity and occurs primarily for colors in the magenta region.

The present study has considered transmission of only the lower bands of coefficients, along with the Band 1 "B" functions for the I and Q chrominance axes. Specifically considered have been the following systems:

System	I	Q
A	Band 3	Band 2
B	Band 2	Band 2
C	Band 2	Band 1
D	Band 1	Band 1 "B" functions only.

The Band number indicates the maximum band for which non-zero value coefficients are transmitted. Coefficients in higher bands are not transmitted and at the receiver are assumed to be zero.

System A is roughly representative of the N.T.S.C. specification relative to horizontal frequency components although the image display device used (a typical device) does not display any chrominance components higher than the Band 2 of the Triangle and Pyramid Transform.

SPECIFIC WORK OF THE DARPA TASK

The Avelex work has been first to construct necessary hardware and software to convert a pre-existing breadboard for monochrome image compression to that capable of processing and compressing full color motion imagery. To accomplish this the previously existing breadboard Pyramid Transformer has been time multiplexed so as to transform, in the language of the N.T.S.C. Video specification, the Y (monochrome), and the chrominance "I"

and "Q" signals in sequence at the transmitter. This accomplishes the complete transformation of all three separate video components into Pyramid transform coefficients. Several new breadboard components including frame stores have been added and/or constructed to do this. Second, software for a personal computer has been written and implemented to control via simple keyboard entry all of the parameters relative to the research of this task. These include the Band number, which controls the frequency groups used in compression of the chrominance components, and the map formation method which determines whether the monochrome signalling map is used for signalling the presence of non-zero value chrominance coefficients or whether individual chrominance maps are formed for such signalling. Also, combinations are possible such that the "Y" and "I" maps can be formed and the "Q" not formed but assumed to be adequately represented by the "Y" map.

Figure 4 shows the modifications of the Pyramid system breadboard to accomplish the addition of the chrominance components. The video Analog-to-Digital converter and digital comb filter were pre-existing to separate the monochrome and chrominance subcarrier signal components from the encoded video input. A new chrominance digital to analog converter has been added followed by an analog chrominance demodulation system such as to recover the "I" and "Q" baseband video signals. Filters in the demodulator limit the "I" bandwidth to 1.5 MHz. and "Q" bandwidth to 0.5 MHz. The "I" and "Q" baseband sigals are then time multiplexed and converted to digital video signals in a six bit Analog-to-Digital converter. New frame stores receive the digitized "I" and "Q" signals to store them until they can be processed in sequence in the single forward Pyramid Transformer. The "Y" signal is delivered to the Transformer directly without need of a frame store but the "I" and "Q" signals are saved over from the same frame that was used for capture of the "Y" signal. The electronic switch ahead of the transformer allows the forward Pyramid Transformer and forward Coefficient Processor, previously implemented, to process in sequence the "Y", "I" and "Q" video frames from the same video input frame.

At the receiver the same previous Reconstruction Coefficient Processor and Pyramid Reconstruction Transformer are used in sequence to reconstruct the "Y", "I" and "Q" video signals and place them in frame stores to be subsequently displayed by the receiver. Two frame stores are used for each of the three signals in a ping-pong fashion to permit simultaneous switching of the same Y-I-Q video frame combination to the display device at one time. That is, while one trio of frame stores is serving to refresh the receiver display, the second trio is receiving sequentially the next "Y", "I" and "Q" video components. A new Digital-to-Analog converter trio and a N.T.S.C. encoder to generate a composite color video signal for output display have been constructed.

CALCULATED CHROMINANCE PYRAMID TRANSFORM EFFICIENCY

In a previous section on the explanation of the Triangle and Pyramid Transforms the relationship between the bands of the transform and the frequency spectra they occupy was established. It was shown that the coefficients of the top band of the transform occupied primarily the top half, or octave, of the frequency spectrum; the next lower band coefficients occupied primarily the next lower octave which is one half the width in frequency of the top octave, and so forth. The proposal for this DARPA contract postulated that it should be possible to omit from transmission the coefficients of one or more of the higher bands of the Pyramid Transform since these bands correspond to spatial frequency components higher than those usually observable by the human eye. At the receiver, the omitted coefficients are given the value of zero and the reconstruction transform subsequently performed.

The efficiency to be gained by omission of the higher band Pyramid chrominance coefficients can be determined mathematically. The total number of transform elements developed for a monochrome image transformed by the Pyramid Transform into five bands is:

$$E = B * (1 + 3*(1 + KO + KO^2 + KO^3 + KO^4)).$$

where B is the number of Band 1 "B" functions and KO=4 for the case with no compression. For a typical B=240 (16 horizontal by 15 vertical), this yields 245,760 elements. Due to the ability of the Pyramid Transform to remove redundancy and hence produce many zero value coefficients which do not require transmission the equation can be modified to include empirically observed factors "A" and a smaller "KO":

$$W = B * (1 + 3*A*(1 + KO + KO^2 + KO^3 + KO^4)).$$

Herein it is assumed that all values of the Band 1 "B" functions require transmission and that "A" represents the fraction of Band 1 coefficients which are typically non-zero and is observed to be 0.8. KO is about 2 for a five band system. The ratio "R" of transform elements to non-zero value transform elements can be calculated as:

$$R = \frac{B * (1 + 3*(1 + (4) + (4)^2 + (4)^3 + (4)^4)}{B * (1 + 3*0.8*(1 + 2 + (2)^2 + (2)^3 + (2)^4)}, \text{ or,}$$

$$R = \frac{1 + 3*(1 + 4 + 16 + 64 + 256)}{1 + 2.4*(1 + 2 + 4 + 8 + 16)}, \text{ or,}$$

$$R = 13.58.$$

The value of 13.58 will be used as a typical value for single image monochrome compression prior to any additional compression obtained by variable length (Huffman) coding of the transform coefficients.

The following calculations assume that all coefficients in a particular band are omitted from transmission when it is determined to exclude coefficient elements from that band. The result is that both vertical and horizontal resolution are decreased by the same amount. This is in contrast to the N.T.S.C. practice of decreasing only the horizontal resolution. The justification for decreasing the resolution in both directions is that the human observer has no better chrominance acuity in one direction than another.

Although calculation results could be herein reported for increased efficiency due only to elimination of higher band chrominance data to be transmitted and not to other efficiencies due to transform compression, such numbers do not reflect the actual chrominance to monochrome ratios in practice. It should also be shown that there are monochrome efficiencies, and that efficiencies accrue more to redundancy removal in the higher bands - the very ones being eliminated in the present chrominance case.

The first case concerns "I" coefficients up through Band 3 and "Q" coefficients through Band 2. The relative amount of data resulting, relative to the monochrome operation, is:

$$1 + \frac{\text{"I" data + "Q" data}}{\text{Monochrome data}} ,$$

and the fraction of data in excess of the monochrome data is:

$$F = \frac{\text{"I" data + "Q" data}}{\text{Monochrome data}} \text{, or,}$$

$$F = \frac{(1 + 2.4*(1 + 2 + 4)) + (1 + 2.4*(1 + 2))}{1 + 2.4*(1 + 2 + 4 + 8 + 16)} ,$$

$$F = \frac{17.8 + 8.2}{75.4}$$

$F = 0.3448$, or $F = 34\%$.

The results of similar calculations for this and other combinations of "I" and "Q" resolution relative to "Y" resolution are shown in Table 1.

Table 1

"I" Band	"Q" Band	"Y" Band	F (Overhead for chrominance)
3	2	5	34 %
2	2	5	22 %
2	1	5	15 %
1	0 (Note)	5	6 %
3	2	4	70 %
2	2	4	44 %
2	1	4	31 %
1	0 (Note)	4	12 %

Note: a "0" indicates that only Band 1 "B" functions are used.

The Table also shows the amount of overhead when the monochrome Band 5 coefficients are omitted from transmission. Again, relative efficiencies have been calculated without regard to additional efficiency gained for both monochrome and chrominance components through use of variable length (Huffman) coding of the transform coefficients.

USE OF MONOCHROME MAP FOR CHROMINANCE COMPONENTS

The second part of the task has been to experimentally evaluate use of a common signalling map to signal the transmission of non-zero value chrominance transform coefficients as well as the non-zero value monochrome coefficients. The map used for monochrome signalling is shown in Figure 18 of Ref 1, and is a tree structure method of signalling from the transmitter to the receiver where to place the sub-set of non-zero value transform coefficients also transmitted for use in the transform reconstruction process. The map is a relatively small amount of overhead compared with the transform coefficient data transmitted and is a very small amount of data compared with all of the zero value coefficients which, as a result of the use of the map, do not require transmission.

It was postulated in the proposal for this task that the same map which is developed, based on the transform coefficient data, for the monochrome, or "Y" component of the image could be used directly for the two chrominance components. Since the first part of this task explored use of only the lower bands of the Pyramid Transform for transmission of chrominance coefficients, the monochrome map would be used for chrominance transform coefficients only in those applicable lower bands.

A color video signal usually is initially generated by a camera which divides light coming through its lens into Red, Green and Blue primary color separations, these being the three additive primaries used for color television imagery. The Red, Green and

Blue (R, G and B) signals are taken as three independent variables since any combination of them is possible at any spatial location within an image. The monochrome "Y" signal and the two chrominance signals, "I" and "Q" are formed as different linear combinations of the R, G and B signals by,

$$Y = 0.299*R + 0.587*G + 0.114*B$$
$$I = 0.596*R - 0.274*G - 0.322*B$$
$$Q = 0.211*R - 0.523*G + 0.312*B .$$

The "Y" signal represents the luminance value of the color signal. Both the "I" and "Q" signals have zero amplitude for the R=G=B condition corresponding to any non-colored portion of an image. The Y, I and Q signals are used in the present task of developing transform coefficients for transmission. At the receiver an inverse matrix of Y, I and Q signals yields the R, G and B signals:

$$R = 1.0*Y + 0.956*I + 0.621*Q$$
$$G = 1.0*Y - 0.272*I - 0.647*Q$$
$$B = 1.0*Y - 1.106*I + 1.703*Q$$

Due to the independence between the R, G and B, and between the Y, I and Q signals there is nothing from the algebra to suggest any commonality between them which would indicate redundency and hence the opportunity for compression. However, the R, G and B signals in practice are very highly correlated with each other when originating from natural scenes. Ref 2, Figures 7-11 and 7-12 between pages 194 and 195 show a color image with accompanying photographs of R, G and B, and Y, I and Q separations wherein the high degree of correlation is quite observable.

Correlation in the present context really refers to the similarity of the spatial location of changes between the Y, I and Q signals. Due to the nature of the Pyramid Transform and its finite and generally narrow footprint on the image signal this "change" correlation in the R, G and B signal domain carries over well into the transform coefficient domain. This is in contrast to most other transformations wherein each of several coefficients is usually a function of all of the same points in the original signal space. In this latter case considerably more points are usually weighted and summed over the same signal space to calculate each of the coefficients, resulting in a loss of correlation between a particular coefficient and a spatial location. Relative to the Pyramid Transform and more specifically, the aforementioned changes refer to departures from spatial linearity of the image separations.

It is known that no non-zero value transform coefficients are generated by the Pyramid Transform for image areas where the intensity of the input signal varies linearly with any spatial direction. Also, the system is configured to transmit only

non-zero value coefficients as directed by the signalling map. The map system accompaying the Pyramid Transform monochrome coefficients serves to indicate any non-zero value coefficient regardless of its algebraic sign or absolute value, if above a small threshold which qualifies it for transmission. Since the map signals no information concerning signa polarity or amplitude, a chrominance coefficient at the same loc cion in the transformation process can have an arbitrary polarity and amplitude relative to its monochrome counterpart and its presence be adequately signalled to the receiver by the monochrome map component.

Although the aforementioned correlations are in practice very strong there is no guarantee of their existence 100% of the time. Two cases exist where a mistake can be made. The first kind of mistake occurs where a monochrome but no chrominance change occurs, and a map component generated and transmitted. A zero value chrominance coefficient is transmitted as a result which does not produce a reconstruction transform error but causes a transmission inefficiency to occur. The second case exists where a non-zero value chrominance component occurs but a zero value monochrome component exists. In this situation the chrominance component is not transmitted and a chrominance reconstruction error occurs. Depending upon the amplitude and Band number, the error may or may not be visible in the reconstructed image.

RESULTS AND CONCLUSIONS

Construction of electronics to add chrominance capability to the Avelex Pyramid Transform experimental system has been completed with all systems functioning satisfactorily. Tests employing a standard Color Bar pattern at the analog input to the compression system show all test colors at the reconstructed output to be within hue and saturation specifications as measured with a standard N.T.S.C. Vectorscope.

Part 1 of the demonstration task has been to show and evaluate color images which have been compressed using the Pyramid Transform, have then had certain higher band coefficients of the two chrominance signals "I" and "Q" discarded, and the images then reconstructed from the remaining coefficients. First, the effect of decreasing vertical color resolution as well as horizontal color resolution, compared with the N.T.S.C. practice of only decreasing horizontal color resolution, is completely acceptable in visual appearance to the point of going unnoticed, except in cases of quite careful observations. Situations where a difference in color performance can be noticed between reducing and not reducing vertical color resolution occur where vertically narrow and horizontally wide colored areas exist, as in the case of letters of the alphabet in color. The letter "T", when occupying a very small part of the image area can be taken as an example. In the N.T.S.C. system, the horizontal line forming the top of the letter shows full chromaticity whereas the vertical line forming the stem of the letter shows reduced saturation of the color. The Pyramid Transform reconstructed letter "T" using lower resolution in both horizontal and vertical directions shows reduced color saturation of both parts.

Of concern prior to construction and evaluation of the color processing to effect reduction in resolution of the color signal components was the possibility of visibly objectionable aliasing in the reconstructed image. This could result since the bands of the Pyramid Transform are only moderately frequency selective and failure to limit the frequency range of the input color components commensurate with the highest band used for chrominance transmission could result in aliasing. This effect has previously been observed with the monochrome signal. The color aliasing effect, although it does occur, is fortunately virtually unobservable. To be able to see the aliasing one must view a highly saturated color at very close range to the receiver and be searching for it.

The efficiency of transmission with reduced horizontal and color resolution is very significant. Considering that three signals must be transmitted to effect a reconstructed color image, as opposed to a single signal to effect a monochrome image, a 200% overhead relative to the monochrome signal is necessary to

transmit a color signal without considering any chrominance resolution reduction. It was shown in Table 1 that the chrominance overhead could be reduced to between 6% and 34% for a full 5 Band system with monochrome resolution of 504 H * 480 V. For a 4 Band system with monochrome resolution of 252 H * 240 V resolution the chrominance overhead could be reduced to between 12% and 70%. In both cases the overhead percentage depends on the chrominance resolution transmitted. In general it has not been possible to distinguish the difference by observation of using Band 2 coefficients for the "I" chrominance signal relative to using Band 3 coefficients. Since the "I" signal corresponds to a chromaticity axis to which the eye is more spatially sensitive it would at first seem that greater detail should be observable with this signal than Band 2. However, most color receivers, including the one used for this experiment, do not process any color in the horizontal direction above the frequencies covered by Band 2. Thus one would not really expect to see horizontal color improvement when the Band 3 chrominance is added. Improvement in the vertical direction by adding Band 3 chrominance appears to be small for naturally occurring scenes. For a color system in which Band 2 is used for both "I" and "Q" signals, the chrominance overhead for a Band 5 monochrome system is 22% and for a Band 4 monochrome system is 44%. Chrominance resolution reduction by another factor of two in each spatial direction is still quite satisfactory for teleconferences and other applications where chromaticity does not have to be observed in great detail and yields overhead relative to the monochrome signal of 15% for a Band 5 system and 31% for a Band 4 system. Further chrominance resolution is possible and is not unacceptable. However, general color saturation is observed to be slightly reduced. The overhead drops to 6% and 12% respectively as a result.

Part 2 of the task has been to evaluate the ability of the monochrome ("Y") signalling map to adequately be used by the two chrominance signals to also signal their non-zero value coefficients. The experiments show, by observation, that signalling of the "Q" component is satisfactorily signalled by the monochrome map for all of the various chrominance resolution reduction systems tried. The results for the "I" component, however, show an occasional observable chrominance error by using the monochrome signalling map. The errors do not appear when using the "I" map directly so the errors seem directly traceable to the monochrome map. The observed errors are sufficiently large in spatial extent to conclude that they probably occur in Band 1 coefficients rather than Band 2 or higher coefficients. Thus the correlation between chrominance changes and monochrome ones is very good, but not 100% for the "I" signal.

Another method to cause the "I" signal to be completely signalled without actually transmitting a separate "I" map is postulated and should solve the observed problem. The method is to produce a logical "OR" of the "Y" and "I" maps, at least for

Band 1 and to transmit that instead of the "Y" map. Transmission of all non-zero values of both the monochrome and "I" signals is thus guaranteed. A small amount of inefficiency results in that some zero value monochrome coefficients are transmitted. However, this is judged preferable to the transmission of a completely separate "I" signalling map. The constructed hardware did not permit this case to be experimentally tried and evaluated.

The chrominance Band 1 "B" functions form the foundation for the color in a particular Pyramid Transform reconstructed image and are adequate to reproduce the correct coloring in all areas of an image, although the color resolution is less than desired. These Band 1 "B" functions can always be transmitted even at very low transmission channel capacities and with continuous motion in the entire image since there are only 240 total "I" and 240 total "Q" of these elements in a color image frame. Therefore no image break-up can occur in the chrominance (or the luminance) portion of the images as the basic Band 1 "B" function building blocks are always given first priority and transmitted without delay. This stability is shown to occur in practice with the experimental hardware. As transmission channel capacity is or becomes available, the chrominance coefficients in Band 1 and those higher bands which are used for color can be transmitted to provide the desired chrominance resolution.

The efficiency of chrominance coding has been shown to be dependent upon the degree of resolution used in a particular system but to be substantially increased by not transmitting resolution which cannot be seen by an observer at the receiver. The efficiency is also increased by using an already existing monochrome signalling map, or one perhaps slightly modified from it to signal the non-zero value chrominance coefficients rather than requiring an extra and independent map for each of the two chrominance signals. The desired decrease in chrominance resolution to achieve the desired efficiency is obtained by simply discarding coefficients in certain higher bands produced by the forward Pyramid Transformer and does not require any additional special hardware.

APPLICATIONS

The Video compression technology developed under this task and previously developed by Avelex provides a combination of performance and simple implementation such that portable, miniaturized, battery powered devices can be a reality for applications including battlefield communications, covert operations, remote sensor and video guidance systems.

The video compression performance of the Avelex Codec has been demonstrated by this task and through previous Avelex work. The technology behind the performance requires hardware which need not perform multiplications or other high power drain operations,

but only additions, subtractions and binary shifts and work at only modest calculation speeds. As a result it is possible, although not yet done, to implement a Transform processor in a low power complementary metal-oxide semiconductor (CMOS) integrated circuit. Also possible because of modest interframe compression processing speed (5 MHz.) is implementation of the control memory as well as RAM storage in CMOS programmable memory. No advances in semiconductor technology are required for this implementation. Although some bipolar circuitry, such as the A/D converter, will be required this circuitry can be duty-cycle power switched to consume power only when actually performing its function and powered down when not required. In the example of the A/D converter, only one out of every eight fields may be processed in a motion image compresion system. The A/D converter need only be operated in that field in which it must perform its conversion. A flash converter unit which usually draws 1.25 watts when used continuously requires less than 160 milliwatts in the above example. A total power consumption for a Pyramid Transform compression system of 20 watts can be realized.

The circuitry which provides the low power in a Transform processor also provides great size reduction. Whereas a video compression system now requires three or four ubic feet and some requiring four times that much, it is possible to implement a Pyramid Transform compression system with silicon integrated transformers in the size of a briefcase (about 500 square inches).

References:

1) U.S. Patent 4,447,886. Triangle and Pyramid Signal Transforms and Apparatus, May 8, 1984, G. William Meeker.

2) Color Television Engineering, John W. Wentworth, McGraw Hill, 1955.

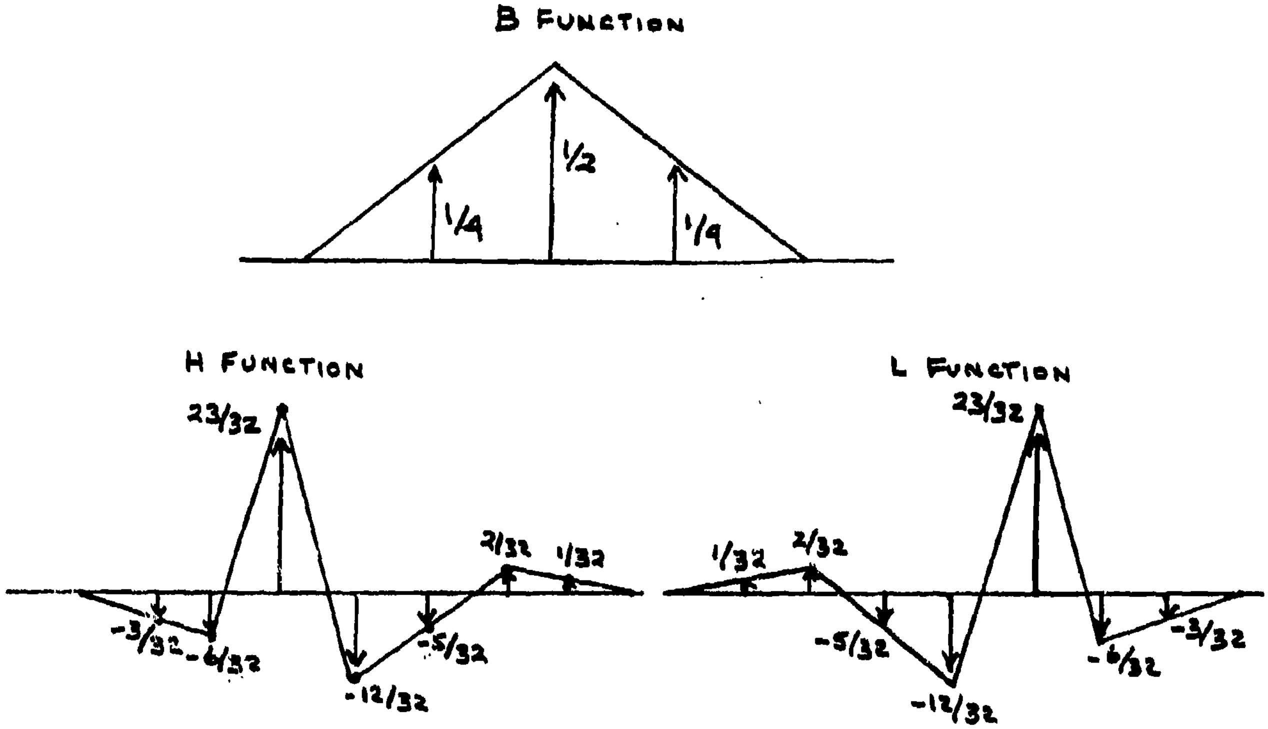

B FUNCTION
1/4
1/2
H FUNCTION
23/32
2/32
1/32
-3/32
-6/32
-12/32
-5/32
L FUNCTION
1/32
2/32
23/32
-5/32
-12/32
-6/32
-3/32

FIG 1

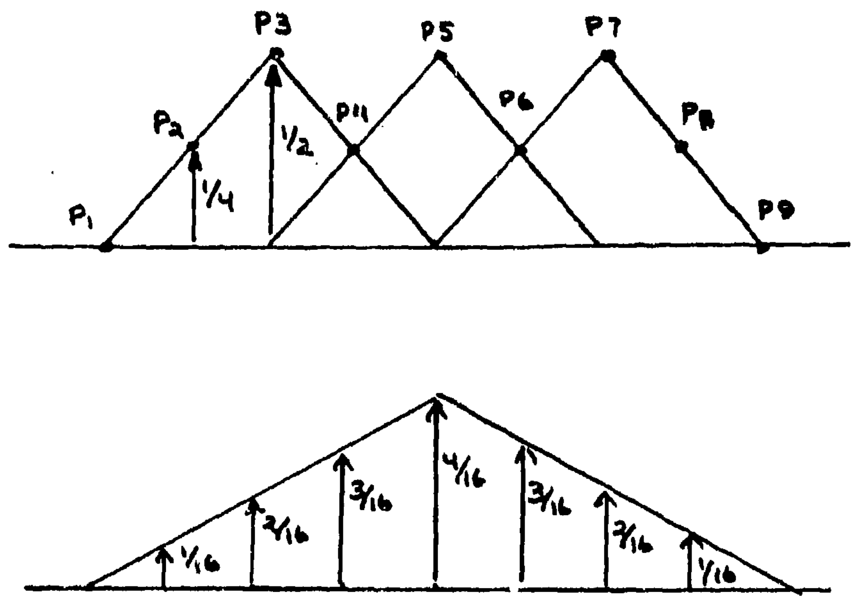
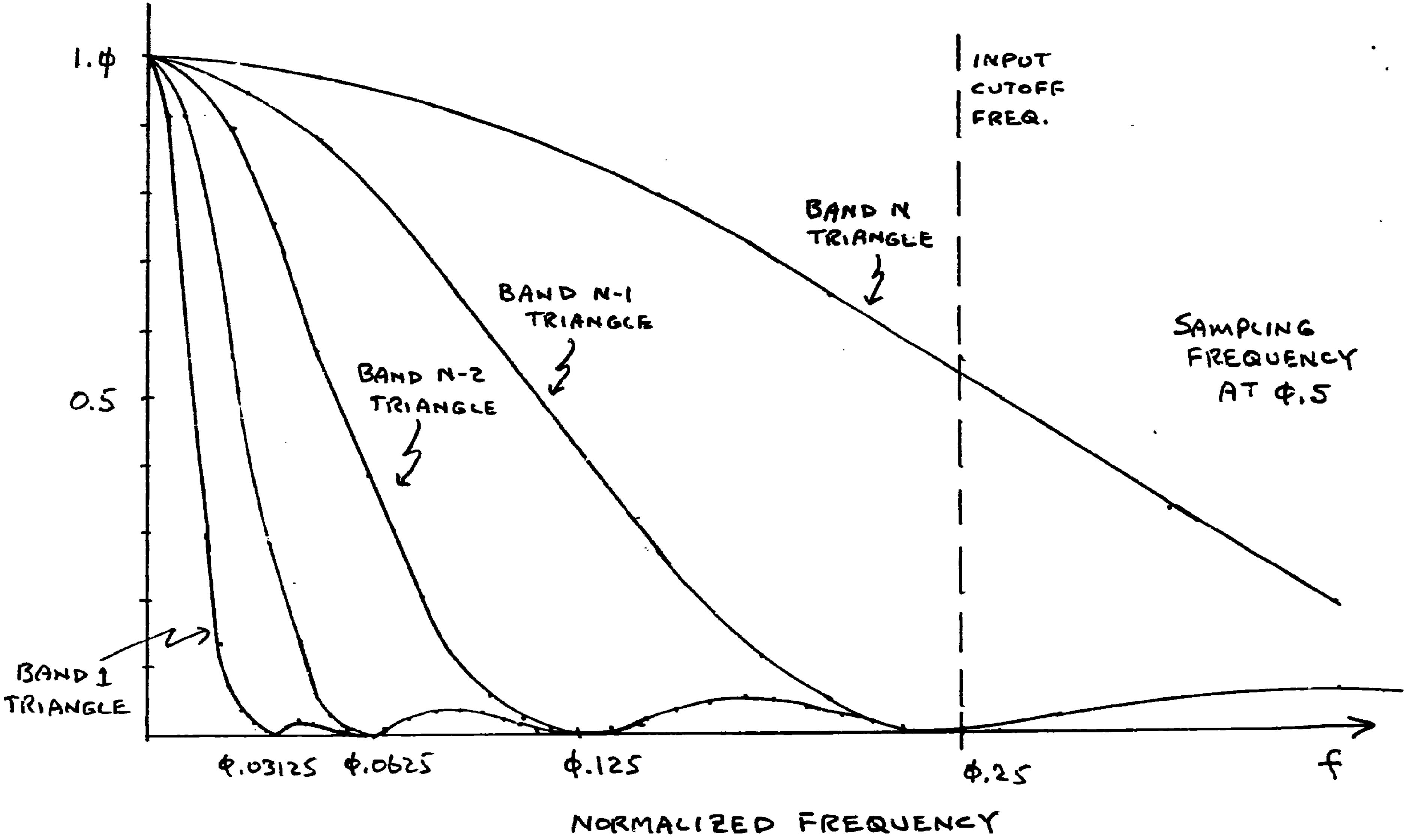

P3
P5
P7
P2
P6
P8
P1
P9
1/2
1/4
1/16
2/16
3/16
4/16
3/16
2/16
1/16

FIG 2



1.Ø
0.5
INPUT CUTOFF FREQ.
BAND N TRIANGLE
BAND N-1 TRIANGLE
BAND N-2 TRIANGLE
BAND 1 TRIANGLE
SAMPLING FREQUENCY AT Ø.5
Ø.03125
Ø.0625
Ø.125
Ø.25
f
NORMALIZED FREQUENCY

FIG 3

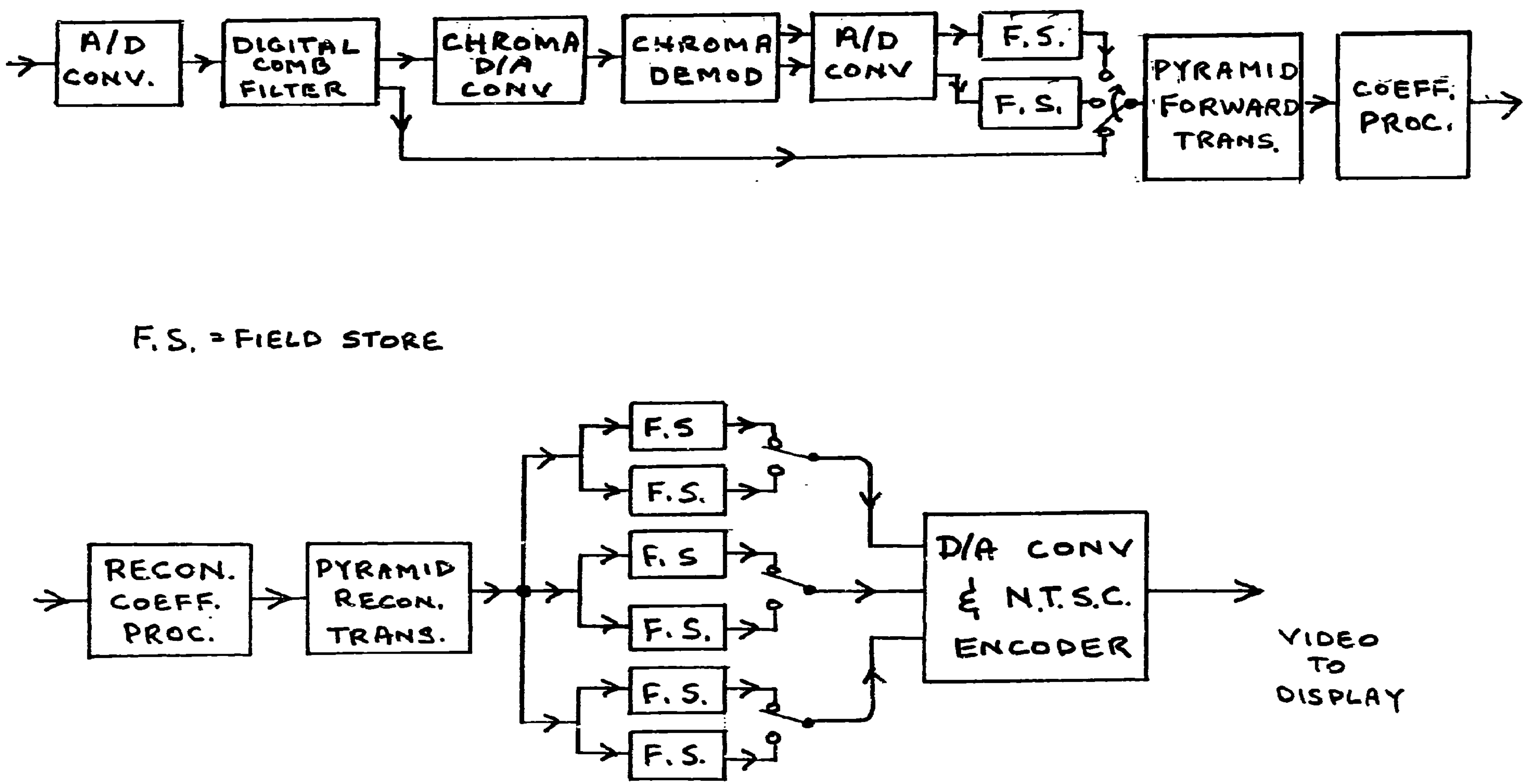

A/D CONV.
DIGITAL COMB FILTER
CHROMA D/A CONV
CHROMA DEMOD
A/D CONV
F.S.
F.S.
PYRAMID FORWARD TRANS.
COEFF. PROC.
F.S. = FIELD STORE
RECON. COEFF. PROC.
PYRAMID RECON. TRANS.
F.S
F.S.
F.S
F.S.
F.S.
F.S.
D/A CONV & N.T.S.C. ENCODER
VIDEO TO DISPLAY

FIG 4